START: Single-to-double Arm Transition Design for Phase II Clinical Trials

Haolun Shi, Teng Zhang, Guosheng Yin
University of Hong Kong

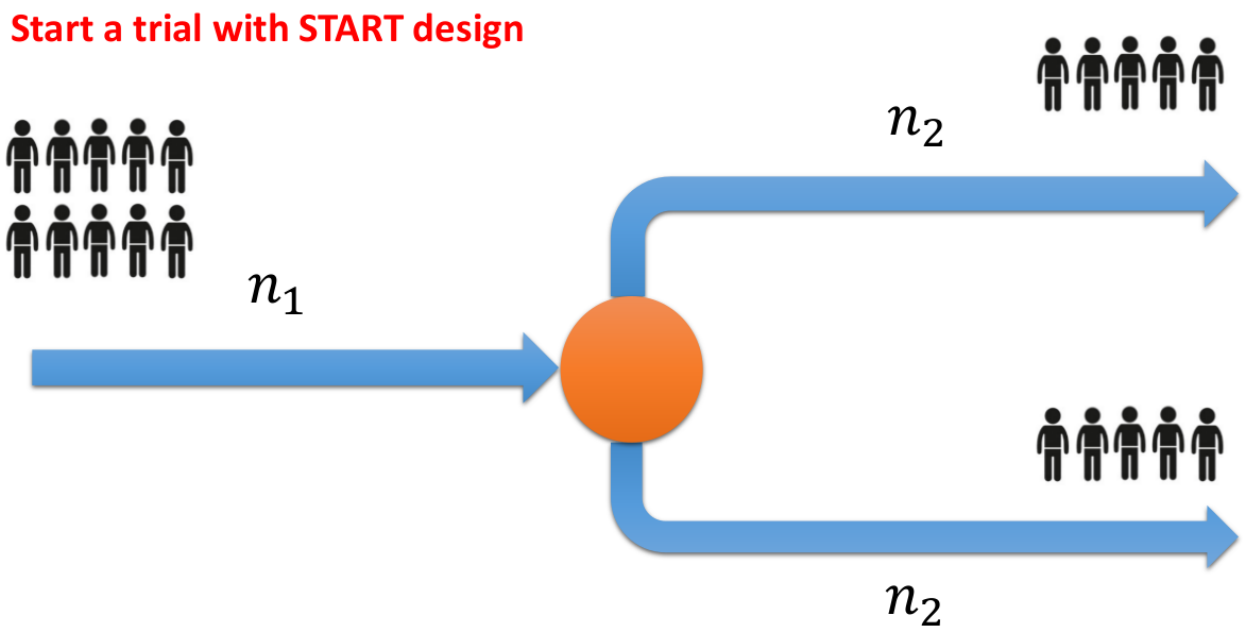
Introduction

- **Phase II clinical trials designed for evaluating a drug's treatment effect can be either single-arm or double-arm.**
- **A single-arm design tests the null hypothesis that the response rate of a new drug is lower than a fixed threshold.**
- **A double-arm scheme takes a more objective comparison of the response rate between the new treatment and the standard of care through randomization.**

Introduction

- **Although the randomized design is the gold standard for efficacy assessment, various situations may arise where a single-arm pilot study prior to a randomized trial is necessary.**
- **To combine the single- and double-arm phases and pool the information together for better decision making, we propose a Single-To-double ARm Transition design (START) with switching hypotheses tests.**
- **The first stage compares the new drug's response rate with a minimum required level and imposes a continuation criterion, and the second stage utilizes randomization to determine the treatment's superiority.**

Trial Illustration.



Design: First Stage

- **In the first single-arm stage, we compare the response rate of the experimental drug with a fixed null rate. The null and alternative hypotheses are formulated as**

$$H_0 : p_E \leq p_0 \quad \text{versus} \quad H_1 : p_E \geq p_1,$$

where p_E is the response rate of the experimental drug, p_0 is the minimally required level for the response rate to be clinically meaningful, and p_1 is the desirable target rate.

- **In the first stage, the trial enrolls n_1 patients, and let x_1 denote the number of responses, which follows a binomial distribution, $x_1 \sim \text{Bin}(n_1, p_E)$. Let r_1 denote the minimum required number of responses for the trial to proceed into the second stage. At the end of the first stage, the decision rules are described as follows:**

  (1) If $x_1 < r_1$, the trial would be terminated early for futility.

  (2) If $x_1 \geq r_1$, the trial would proceed into the second stage, where a total number of $2n_2$ patients are equally allocated to the experimental and standard arms.

Design: Second Stage

- **In the second stage, which takes a double-arm randomization scheme, we assess the superiority of the experimental drug in a comparison with the standard treatment. The testing hypotheses are switched to**

$$H_0 : p_E \leq p_S \quad \text{versus} \quad H_1 : p_E > p_S,$$

where p_S is the response rate of the standard treatment. Let x_2 and y_2 denote the numbers of responses in the experimental arm and the standard arm, respectively.

Design: Second Stage

- **At the end of the trial with completion of both stages, the decision rules are based on the Z-test statistic,**

$$T_Z = \frac{\hat{p}_E - \hat{p}_S}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1+n_2} + \frac{1}{n_2}\right)}},$$

where $\hat{p}_E = (x_1 + x_2)/(n_1 + n_2)$, $\hat{p}_S = y_2/n_2$, and $\hat{p} = (x_1 + x_2 + y_2)/(n_1 + 2n_2)$ are the sample proportions in the experimental arm, the standard arm, and both arms combined, respectively.

- **At the end of the second stage, the decision rules are described as follows:**

  (1) If $T_Z > c$, where c denotes the critical value for the Z-test, we declare the drug as promising.

  (2) Otherwise, we declare the drug nonpromising.

Design: Error Rates

- **Let α_1 and β_1 denote the frequentist type I and type II error rates for the first stage, respectively. Based on the decision rules in stage 1, we have**

$$\alpha_1 = \sum_{x_1=r_1}^{n_1} P(x_1|p_0) = 1 - F_{\text{Bin}}(r_1 - 1; n_1, p_0),$$

$$\beta_1 = \sum_{x_1=0}^{r_1-1} P(x_1|p_1) = F_{\text{Bin}}(r_1 - 1; n_1, p_1),$$

where F_{Bin} denotes the cumulative distribution function for the binomial distribution.

Design: Second Stage

- **Let α_2 and β_2 denote the frequentist type I and type II error rates for the second stage, respectively.**
- **In the second stage, we are testing the hypotheses $H_0 : p_E \leq p_S$ versus $H_1 : p_E > p_S$. To control the type I error rate, we consider the scenario $p_E = p_S = p_0$, under which the probability of rejecting the null should be below α_2.**
- **Similarly, to control the type II error rate, we consider the scenario where there is a treatment difference between the two arms, $p_S = p_0$ and $p_E = p_1$, and compute the probability of the trial continuing into the second stage and fail to reject the null, which should be controlled below β_2.**

Following the rationale of the trial conduct, we have

$$\alpha_2 = \sum_{x_1=r_1}^{n_1} \sum_{x_2=0}^{n_2} \sum_{y_2=0}^{n_2} P(x_1|p_0)P(x_2|p_0)P(y_2|p_0)I(T_Z \geq c),$$

$$\beta_2 = F_{\text{Bin}}(r_1 - 1; n_1, p_1) + \sum_{x_1=r_1}^{n_1} \sum_{x_2=0}^{n_2} \sum_{y_2=0}^{n_2} P(x_1|p_1)P(x_2|p_1)P(y_2|p_0)I(T_Z < c),$$

where $I(\cdot)$ is the indicator function.

Design

- **Because the implementation of the Z-test in the second stage is contingent upon the decision of continuation at the end of the first stage, the type I and type II error rates across the two stages satisfy $\alpha_1 > \alpha_2$ and $\beta_1 < \beta_2$, respectively.**
- **To maintain the type I error rate for the second stage at α, we need to set $c = z_\alpha$, where z_α denotes the $100(1-\alpha)$th percentile of the standard normal distribution.**

Design

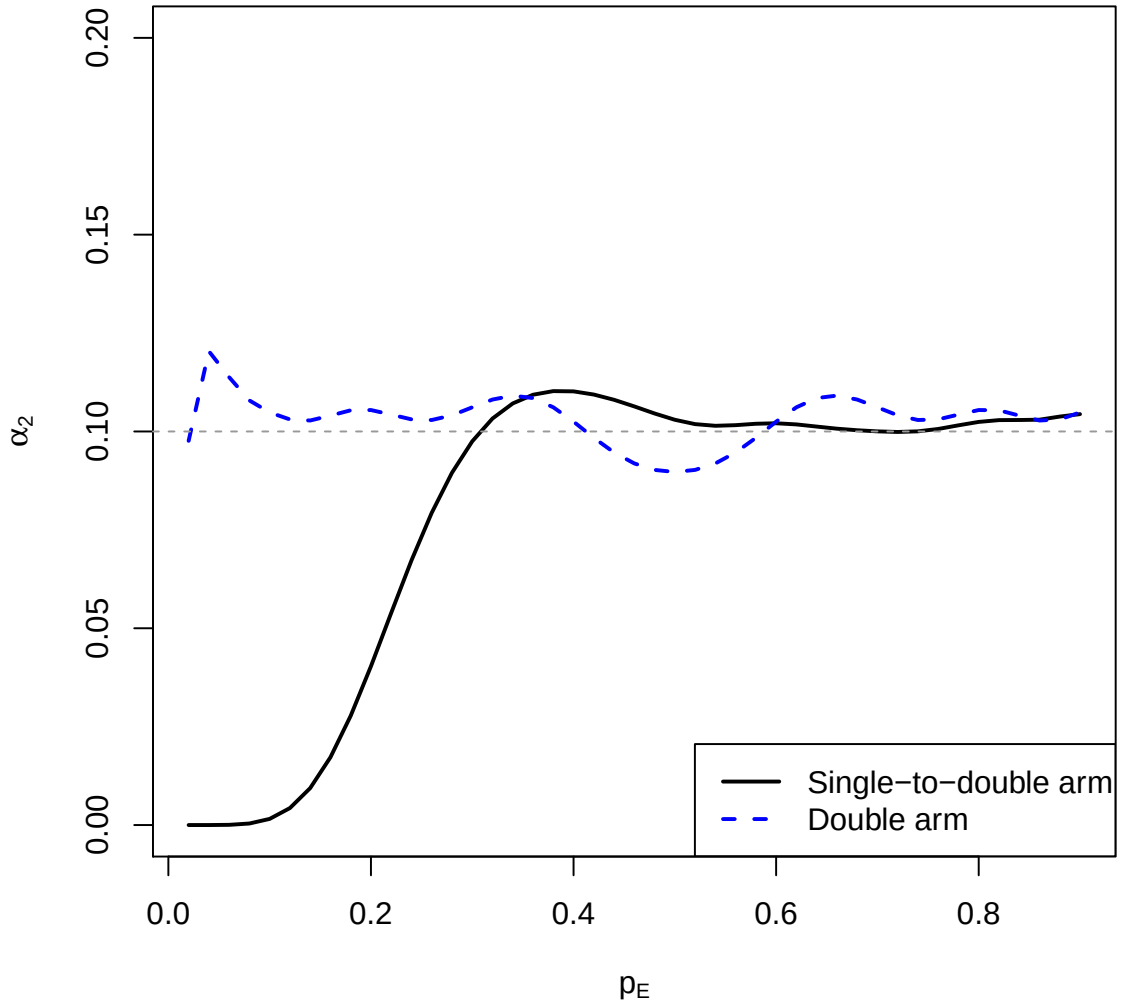
- **To calibrate the design parameters (n_1, n_2, r_1), we need to specify p_0, p_1, and the type I and type II error rate constraints in both stages $(\alpha_1, \alpha_2, \beta_1, \beta_2)$. The optimal set of (n_1, n_2, r_1) is chosen by minimizing the average sample number (ASN), which is the average of the expected sample size under the null (ESS_0) and that under the alternative (ESS_1) hypotheses, ASN= $(\mathrm{ESS}_0 + \mathrm{ESS}_1)/2$, with**

$$\mathrm{ESS}_0 = n_1 + 2n_2\, P(x_1 \geq r_1 | p_E = p_0),$$
$$\mathrm{ESS}_1 = n_1 + 2n_2\, P(x_1 \geq r_1 | p_E = p_1).$$

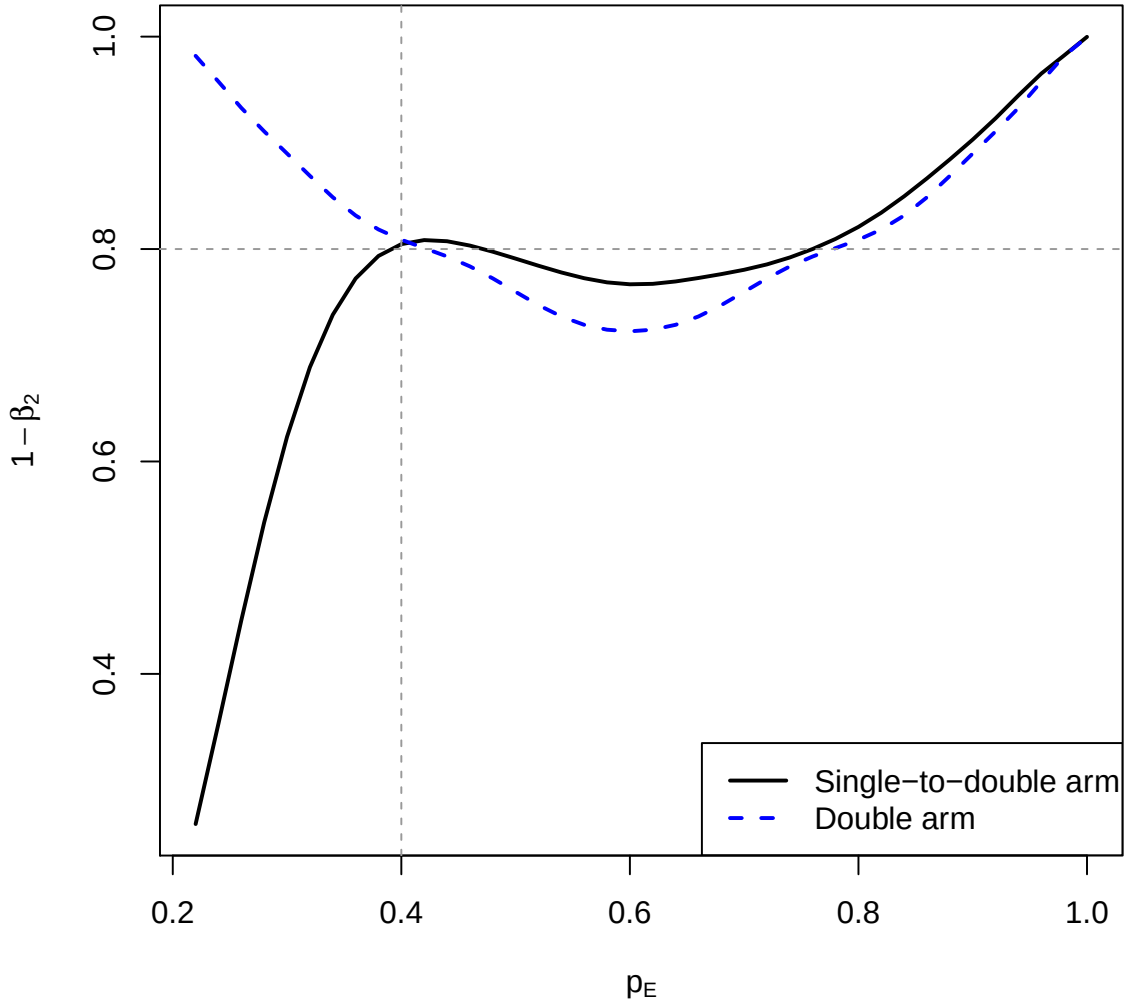
p_0	p_1	Design	n_1	n_2	r_1	$P(\text{S2}\|p_0, p_0)$	$P(\text{S2}\|p_1, p_1)$	$P(\text{S2}\|p_0, p_1)$	ESS_0	ESS_1	ASN
0.1	0.3	IIa+IIb	21	44	3	0.016	0.094	0.803	34.4	101.5	67.9
		START	23	29	3	0.042	0.108	0.800	34.2	77.9	56.0
		Phase IIb	0	35	0	0.105	0.098	0.805	70.0	70.0	70.0
		Simon	5	9	1	0.127	0.900	0.900	8.7	13.3	11.0
0.2	0.4	IIa+IIb	27	62	7	0.016	0.092	0.800	46.3	139.2	92.7
		START	32	38	8	0.040	0.110	0.805	45.3	103.6	74.5
		Phase IIb	0	46	0	0.102	0.107	0.809	92.0	92.0	92.0
		Simon	5	7	1	0.195	0.911	0.911	9.7	11.8	10.7

Design comparison. Phase IIb design is a fully randomized double-arm design, phase IIa+IIb design is a concatenation of separate single-arm and double-arm design. The abbreviation S2 represents the trial success in the second stage.

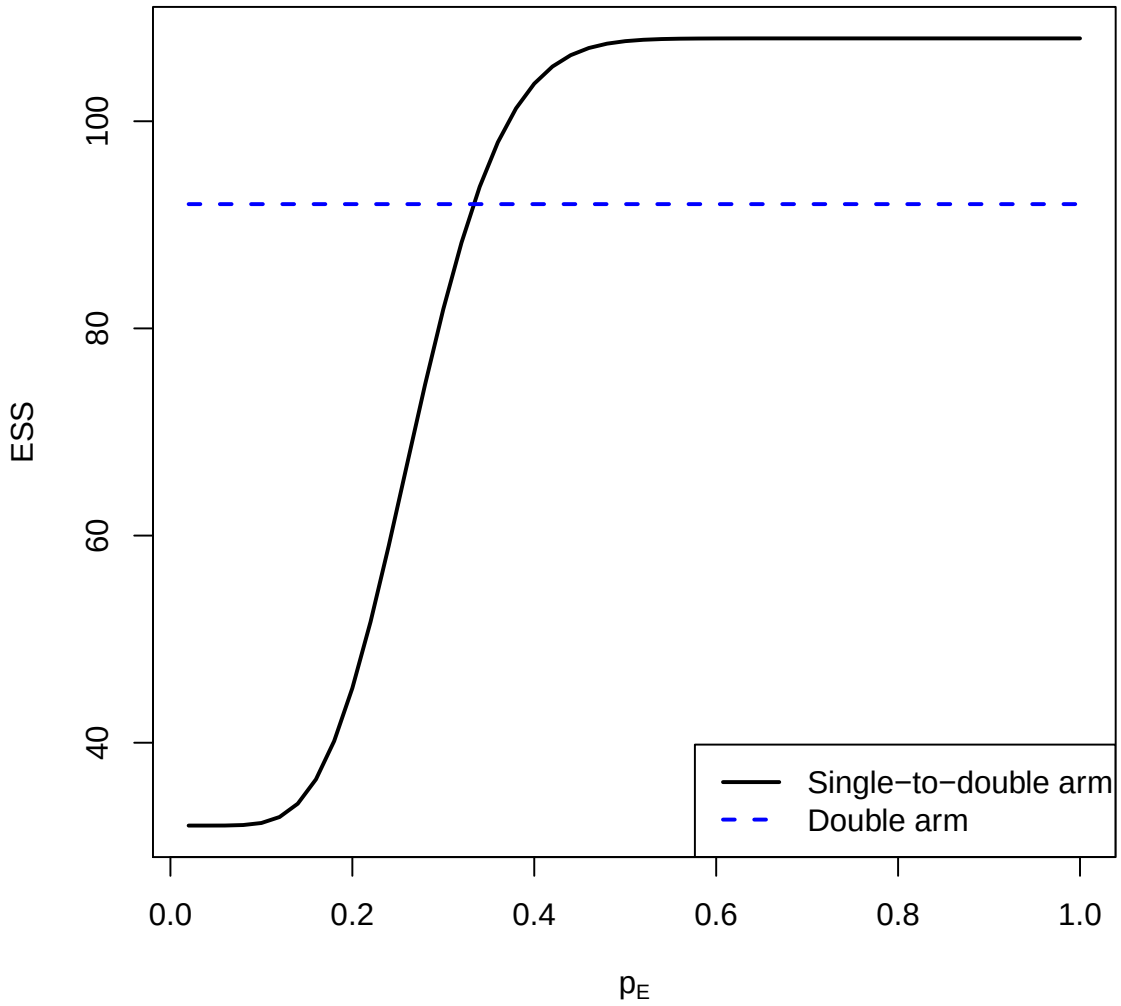
Type I error rate. Conservative due to the existence of the threshold r_1 at the end of the first stage.



Design power. Power reduction due to the existence of the threshold r_1 at the end of the first stage.



Expected sample size.



Discussion

- **Since patients from both the first and second stages are pooled together for final decisions, investigators of the trial need to be precautious about the potential bias incurred during the patient accrual in the single-arm stage, as patients who did not respond to standard treatment might be more enthusiastic to be enrolled in the single-arm stage. It is crucial for the investigator to seek to ensure a balanced allocation of patients' prognostic factors not only in two arms, but also across the two stages. Blinding the investigator and trial participants from knowing the existence of the single-arm stage may mitigate the bias.**

- **Moreover, it is worth noting that at the end of the first stage of the START design, the patient enrollment would halt until the last patient's outcome is observed and the test statistic is computed. Therefore, the START design is more suited for outcomes that can be quickly ascertained after receiving the treatment, otherwise, the gap time between the first and the second stages would be too long for the START design to render any advantage in efficiency over the conventional separate phase IIa plus IIb design.**

Thank you for attending!